FEBS Letters 584 (2010) 4401–4407

journal homepage: www.FEBSLetters.org

Regulation of splicing enhancer activities by RNA secondary structures

Wei Liu [a], Yu Zhou [a,b,c], Zexi Hu [a], Tao Sun [a], Alain Denise [c,d,e], Xiang-Dong Fu [a,b,*], Yi Zhang [a,f,*]

[a] State Key Laboratory of Virology, College of Life Sciences, Wuhan University, Wuhan, Hubei 430072, China
[b] Department of Cellular and Molecular Medicine, University of California, San Diego, La Jolla, CA 92093-0651, USA
[c] Univ. Paris-Sud, LRI UMR8623 and IGM UMR8621, Orsay F-91405, France
[d] CNRS, Orsay F-91405, France
[e] INRIA, Saclay, F-91400, France
[f] ABLife Inc, Building E-1301, 115 Guangba Road, Wuhan, Hubei 430072, China

ARTICLE INFO

Article history:
Received 28 August 2010
Accepted 28 August 2010
Available online 1 October 2010

Edited by Michael Ibba

Keywords:
Alternative splicing
Secondary structure
ESE

ABSTRACT

In this report, we studied the effect of RNA structures on the activity of exonic splicing enhancers on the SMN1 minigene model by engineering known ESEs into different positions of stable hairpins. We found that as short as 7-bp stem is sufficient to abolish the enhancer activity. When placing ESEs in the loop region, AG-rich ESEs are fully active, but a UCG-rich ESE is not because of additional structural constraints. ESEs placed adjacent to the 3′ end of the hairpin structure display high enhancer activity, regardless of their sequence identities. These rules explain the suppression of multiple ESEs by point mutations that result in a stable RNA structure, and provide an additional mechanism for the C6T mutation in SMN2.

1. Introduction

Mammalian genes are commonly interrupted by multiple introns, and the spliceosome-catalyzed intron removal and exon ligation, termed pre-mRNA splicing, represents an important layer of gene regulation [1]. Deep sequencing analysis of the transcriptomes in 15 diverse human tissues and cell lines reveals that 92–94% of human genes are alternatively spliced [2]. Alternative splicing of many mammalian genes are regulated during cell differentiation, proliferation and apoptosis, development, or in response to environmental stimuli [3–6].

Alternative splicing of pre-mRNAs is controlled by precise and highly dynamic interactions between trans-acting regulatory proteins and their corresponding cis-acting splicing regulatory elements (SREs) residing in exons and/or introns. Analysis of such interactions has recently been extended to the genomic level by coupling immunoprecipitation of splicing factor-bound RNA with deep sequencing, termed CLIP-seq or HITS-seq [7].

Although the importance of RNA structures in splicing regulation has been increasingly appreciated [8,11], the prevalence of this layer of regulation is further underscored by a recent genome-wide study suggesting that about 15% of alternative splicing events in the human genome may subject to influence by RNA secondary structures [10]. Most of previous studies focused on the impact of RNA secondary structures on modulating the accessibility of the 5′ and 3′ splice sites to U1 and U2 snRNPs, respectively, or the recognition of the polypyrimidine tract and 3′ splice site by U2AF65-U2AF35 heterodimers. RNA structures that cause sequestration of these essential splicing signals inhibit splicing of associated exons, while other RNA structures that bridging long-distance RNA–RNA interactions in the intron promote splicing by enhancing splice site pairing (lately reviewed by Warf and Berglund [11]).

Besides essential splicing signals (5′ and 3′ splice sites, the pytract and A branch point), exons and introns that are subjected to alternative choices are often enriched with positive and negative SREs, which can be experimentally identified by mutational analysis or predicted by computational algorithms based on existing experimental evidence [7,12,13]. These SREs can be divided into exonic and intronic enhancers (ESE and ISE) and exonic and intronic silencers (ESS and ISS). These regulatory sequences are central to decipher the "splicing code" governing the splicing outcome of most higher eukaryotic genes, which have received considerable attention in the last decade [14–20]. However, although it is conceivable that RNA secondary structures may modulate the accessibility of regulatory proteins to their corresponding exonic and intronic SREs, the progress has been slow. One of the best-understood cases is the regulated accessibility of ESE for SF2/ASF in the EDA exon of the fibronectin pre-mRNA by a fold-back RNA structure, which enhances splicing by presenting the ESE in a

* Corresponding authors. Addresses: ABLife Inc, Building E-1301, 115 Guangba Road, Wuhan, Hubei 430072, China. Fax: +86 27 68754945 (Y. Zhang); Department of Cellular and Molecular Medicine, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093-5004, USA. Fax: +1 858 822 6920 (X.-D. Fu).
E-mail addresses: xdfu@ucsd.edu (X.-D. Fu), yizhang101@hotmail.com (Y. Zhang).

0014-5793/$36.00
doi:10.1016/j.febslet.2010.09.039

single-stranded loop [21,22]. Computational analysis indicates that some known SREs appear to be more single-stranded than control sequences, and consistently, exonic SREs engineered in the loop of a stable hairpin structure display much higher splicing enhancing activities than those in the stem [23].

Here, we further characterized how local RNA structures regulate the activity of ESEs by engineering ESEs into the basepaired stems with varying thermostability, into the loops of varying sizes, or into the single-stranded region associated with a stem to different degrees. Interestingly, in contrast to GA-rich ESEs, an UCG-rich ESE compromises most of its enhancer activity in the loop region. ESEs inserted immediately downstream of the hairpin stem are more active in enhancing the inclusion of the alternative exon. We also present evidence that the enhancer strength of multiple ESEs in an alternative exon is predictable when the effect of local RNA structures is taken into account.

2. Materials and methods

2.1. Plasmid constructs and mutagenesis

The minigene *SMN1* and *SMN2* containing exon 7 (54-bp) and portions of the adjacent intron 6 and 7 from PCI-*SMN1* and PCI-*SMN2* (gifts from Dr. J. Zhou, Arizona State University) was cloned into the GFP coding sequence of pZW8 (a gift from Dr. C. B. Burge, Massachusetts Institute of Technology) to replace the original *SIRT1* minigene, and therefore yielded pZW9-*SMN1* and pZW9-*SMN2*. To insert the desired ESE sequences and hairpin structures, two restriction enzyme sites *Xho*I and *Eco*RI were cloned into exon 7 of pZW9-*SMN1* to replace the Tra2β-ESE sequence (5′ AAAAGAAGGAAGGTG) to create pZW9-*SMN1*-m used in this study.

To make 5D, 5D-L, 5D-a, ASF-L and Tra2β-L series of constructs, sense and antisense oligos containing a full or truncated *Xho*I and *Eco*RI restriction sites were synthesized (SBS Genetech, Beijing) and diluted in the linker buffer (50 mM Tris–HCl pH 8.0, 100 mM NaCl, 1 mM EDTA) at a final concentration of 100 nM. Some oligos contained truncated restriction sites to avoid the splicing-induced frameshift and the consequent non-sense mediated mRNA decay. Each pair of sense and antisense oligos were mixed for annealing: denaturation at 95 °C for 2 min, annealing at 52 °C for 10 min, then chilled on ice. The product diluted to a concentration of 10 nM was then ligated into the *Xho*I and *Eco*RI-digested pZW9-*SMN1*-m.

ASF-a and Tra2β-a constructs were made according to one-step site-directed mutagenesis as previously described [24], using pZW9-*SMN1* as the template and high-fidelity DNA polymerase KOD PLUS (TOYOBO).

For all the construct designs, the unintended ESEs were avoided according to a computational method (manuscript in preparation). The corresponding oligo sequences are listed in the Supplementary data.

2.2. Cell culture and transfection

HeLa (human cervical carcinoma) cells were maintained in Dulbecco's modified Eagle medium supplemented with 10% fetal calf serum. Cells were planted on the 24-well plate with cell concentration of 8×10^4 per well. Cells were transfected when at 60% confluence. The transfection followed standard protocol of Lipofectamine-2000, and 500 ng plasmid was used per well. The cell were kept 24 h after transfected, then harvested for total RNA preparation.

2.3. RNA preparation and RT-PCR analysis

Total RNA was extracted from transfected cells by a one-step extraction method with trizol (Invitrogen). A standard reverse transcription (RT) protocol supplied by Promega (whose MMLV was utilized) was used with oligo$(T)_{18}$, followed by a PCR amplification with the products being separated on 2% agarose gels. The gel images were captured by GeneGenius Gel Imaging System (SynGene) and quantified by the associated GeneTools analysis software.

3. Results

Pre-mRNA folding is limited to a region of about 50-nt downstream of the transcribing polymerase [25], and thus secondary structures of pre-mRNA tend to form locally whereas those require long-range interactions is disfavored [26]. Simple fold-back hairpin structures involving only short-range local basepairing are prevalent in pre-mRNAs, which can be predicted by a number of available algorithms [27]. Although accurate prediction of complex RNA structures remains a main challenge in the field, the local hairpin structures formed within 50-nt of a pre-mRNA can be reliably predicted by minimal free energy algorithms.

3.1. Stems as short as 7-bp block the ESE function

In order to assess the effect of this simple class of RNA secondary structures on the enhancer activity of ESEs, we constructed a *SMN1* (survival of motor neuron 1, telomeric copy) based splicing reporter in which two restriction sites were introduced adjacent to the ESE sequence for the splicing activator Tra2β (Tra2β-ESE) located in the alternative exon 7 (Fig. 1A). This allows insertion of various ESE-containing structures to replace the Tra2β-ESE (Fig. 1A and B). As for formation of the local alternative secondary structures is known to compete with each other and influence their functions [28], we specifically designed various hairpin structures to exclude the possibility of forming competitive alternative secondary structures other than the desired structure within the 50-nt folding window.

Most well-characterized ESEs are AG-rich. However, an unbiased computational analysis revealed that some of the predicted ESEs are CT-rich, as exemplified by Group 5D with a TCGTCG consensus [12]. Fig. 1 shows that a heptamer sequence containing the consensus has a strong ESE activity comparable to that of the Tra2β-ESE, validating the biological importance of this group of ESEs (Fig. 1B and C, Construct 5D). However, the enhancer activity is completely abolished when placing the ESE in the stem region of hairpin structures containing 13–7 contiguous basepairs (Fig. 1B and D, all other constructs except for 5D). This result supports the previous observation that the single-strandedness of SREs is important for efficient recognition by trans-acting splicing factors [23].

3.2. Structural constraints of the hairpin loop interfere with ESE activity

The lost splicing enhancer activity of the ESE-5D series in Fig. 1 is likely due to the hairpin structure that blocks the ESE function by base-pairing. In order to provide further support to this conclusion, we placed two known ESEs, a strong Tra2β-ESE and a weaker ASF-ESE from EPB41 [29], in the loop of a stable hairpin structure (Fig. 2A). We observed either comparable (all Tra2β-ESE constructs and ASF-ESE construct L-3), or higher (ASF-ESE constructs L-0 and L-7) enhancer activities than the parental construct containing a structure-free single-stranded ESE (Lane "SMN1" for Tra2β-ESE and lane "ASF linear" for ASF-ESE) (Fig. 2D and C). This result demonstrated that a simple stable hairpin alone in the exon is not sufficient to have a negative impact on splicing.

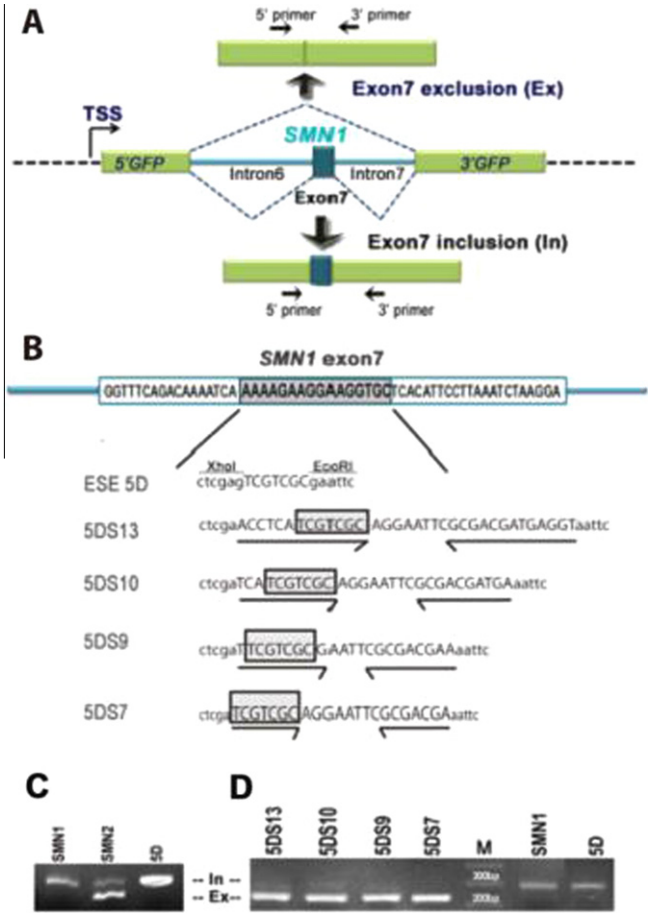

Fig. 1. Inhibition of the ESE function by short stems. (A) Diagram of the splicing reporter pZW9-*SMN1*. The *SMN1* splicing cassette (intron6-exon7-intron7) was inserted to split the *GFP* coding sequence. Transcription of the split *GFP* gene from the indicated TSS (transcription start site) was driven by the CMV promoter. (B) Diagram of the constructs containing an ESE sequence TCGTCGC (referred to as 5D, [12]) in a single-stranded RNA region (ESE 5D) or basepaired stems (5DS series). The basepairs in the stems varying from 13, 10, 9 to 7, resulting in constructs 5DS13, 5DS10, 5DS9 and 5DS7, respectively. The restriction sites used for cloning ESEs were indicated in lower case, which were *Xho*I site upstream and *Eco*RI downstream. The original sequence of *SMN1* exon 7 labelled with gray rectangle is shown on the top, with Tra2β-ESE being shadowed in grey. The engineered sequences replacing the Tra2β-ESE are shown at the bottom. The sequences involving in forming the predicted stem structure are underlined by arrows toward the loop. (C) RT-PCR analysis of wild type *SMN1* and *SMN2* minigenes and construct 5D in transfected HeLa cells. (D) RT-PCR analysis of the 5DS series of constructs containing 13–7-bp stems. Bands labeled by "In" and "Ex" correspond to exon 7 inclusion and exclusion, respectively. Lane M indicates the molecular weight markers.

It has been well-established that nucleotides in some hairpin loops may form non-canonical interactions [30]. For example, UUCG is a highly stable tetraloop invoking non-canonical G-U base pairing and base stacking interactions [31,32]. Such structural constraints would prevent the loop nucleotides from freely interacting with corresponding RNA binding proteins. Therefore, it is theoretically possible that some ESEs in loop regions may experience additional structure interferences, thereby reducing their splicing enhancing activities.

The high enhancer activity of Tra2β-ESE and ASF-ESE in the hairpin loop suggests that these AG-rich ESE sequences are linear and free of structural constraints (Fig. 2B and C). Different from the AG-rich Tra2β-ESE and ASF-ESE, ESEs in group 5D is CU-rich and contains two UCG repeats. These UCG repeats are expected to form non-canonical G–U base pairing and extensive base stacking interactions in hairpin loops. The enhancer activity of ESE-5D in hairpin loop is predicted to be compromised because of these structural constraints. We tested this hypothesis and found that ESE-5D in the hairpin loop (5DL-0) indeed lost most of its enhancer activity (Compare between lane 5DL-0 in Fig. 2D and lane 5D in Fig. 1C). Interestingly, introducing an adenosine to each side of the loop in ESE-5D (5DL-2) strongly increased the enhancer activity, consistent with the possibility that these two adenosine residues disturbed the essential base stacking interactions between the UCG repeats and the first C–G pair in the stem. In contrast, adding U and C in between the A-bracketed UCGUCGC loop would provide additional intra-loop interaction opportunities. As expected, we observed reduced enhancer activity with these constructs (Compare lanes 5DL-5L and 5DL-5R with 5DL-2 in Fig. 2B).

These results together have led to two important conclusions. First, the regulatory activity of SREs in hairpin loops is sequence-specific; the structurally constrained UCG-rich elements are much less active in hairpin loops than the structure-free AG-rich elements. The loop-repression rule may apply to many SREs if they are prone to form some non-canonical interactions in hairpin loops, although the repression strength may vary and be smaller than the stem-repression. Second, spacer nucleotides flanking an ESE in the loop do not appear to engage in interactions with regulatory proteins (Fig. 2B–D). Instead, spacer nucleotides may strengthen or weaken intra-loop interactions to positively or negatively modulate ESE activities.

3.3. ESEs located downstream of a stable hairpin structure are highly active

The splicing regulatory elements may reside either within a local hairpin structure or near a hairpin structure in different pre-mRNAs. It is unclear yet how a hairpin structure might affect a neighboring SRE. In order to address this question, representative ESEs for ASF, Tra2β or belonging to group 5D were each placed immediately downstream of a stable hairpin structure (Fig. 3A). We observed strong enhancer activity with Tra2β-ESE when it was located after the hairpin stem (Fig. 3D). Strikingly, we also detected a high enhancer activity for both ESE-5D and ASF-ESE when placed next to the stem (Fig. 3B and C compared to Fig. 2). The enhancer activity of ESE-5D after the hairpin stem was not fully recovered (5Da-0 in Fig. 3B and ESE-5D in Fig. 1B), which could be due to the nucleotides followed the ESE-5D consensus in the 5Da-0 construct that may create a new splicing silencer (cgtcgcaa), as predicted by Human Splicing Finder [33]. Interestingly, increasing the distance between the hairpin structure and ESE-5D (5Da-1, 5Da-2 and 5Da-3) further reduced the enhancer activity of this splicing regulatory element (Fig. 3A and B), indicating that the regulatory activity of ESE can be maximized when it is right adjacent to the hairpin.

We conclude from these results that hairpin structures immediately upstream of an ESE do not interfere with the activity of SREs. Instead, this structural organization may prevent other sequences from interacting with ESEs to enforce their single-stranded conformation, which may favor their interactions with trans-acting regulatory proteins.

3.4. A stable hairpin structure can interfere with multiple ESEs

The above results suggest that precise evaluation of the activity of ESEs requires the consideration of their structural contexts. The distribution of ESEs in mammalian pre-mRNAs tends to be clustered [34]. Here we attempted to integrate the ESE sequence and structure information to predict the splicing outcome of several engineered sequences containing a cluster of ESEs for SR proteins, namely mulESE-SRs (Fig. 4). Each of the engineered sequences is of 36-nt in length, which was used to replace the Tra2β-ESE (16-nt) in

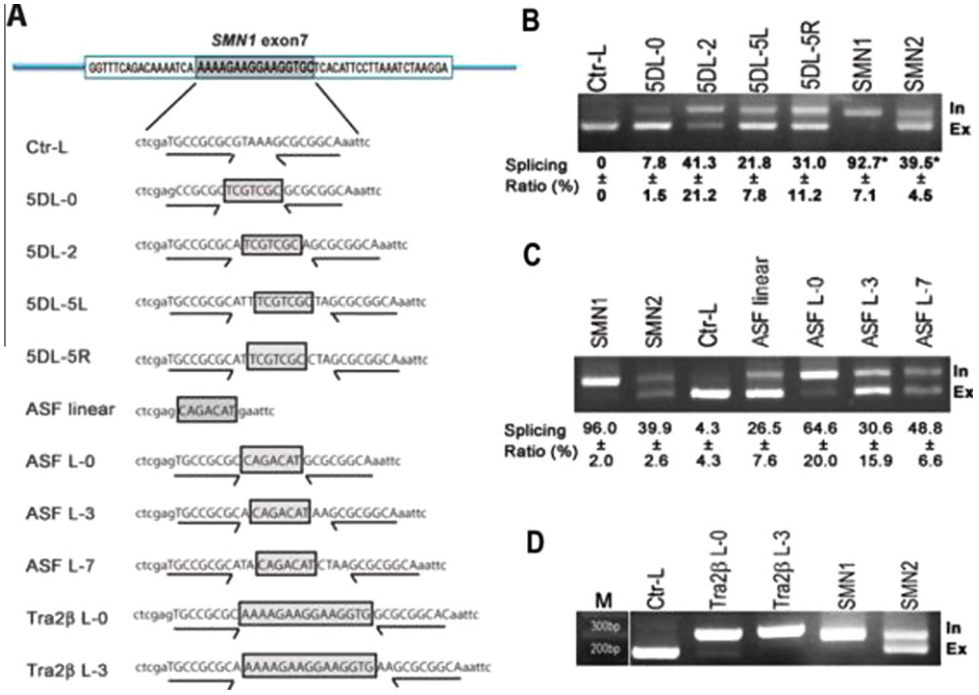

Fig. 2. The enhancer activity of ESEs in hairpin loops depends on their sequence composition. (A) Diagram of the constructs with various ESEs (boxed in grey) engineered into hairpin loops. Ctr-L serves as a control containing a hairpin structure without an ESE. In the "ASF-linear" construct, the ASF-ESE from exon 16 of the EPB41 gene encoding erythrocyte membrane protein 4.1R was cloned into the *Xho*I and *Eco*RI sites in the pZW9-*SMN1* vector. (B–D) RT-PCR analysis of the 5D-L (B), ASF-L (C) and Tra2β-L (D) series of constructs in transfected Hela cells, with the spliced products being resolved on agarose gels. One representative gel for the constructs of each of the three ESEs is shown. Splicing ratios indicate the average percentage of the spliced product with the exon7 inclusion (In) over all spliced products (a sum of those with exon7 inclusion and exclusion, In + Ex), which were calculated from three biological repeats (asterisk indicates the splicing ratios calculated from two biological repeats in a few cases). Experimental variations are indicted by the standard error of the mean (S.E.M) in each case. Constructs with Tra2β-ESE always spliced to full exon7 inclusion (D). Relative large batch variations were observed for exon7 inclusion in constructs containing ESE-5D (B) and ASF-ESE (C), and gel results from the experiments displaying the most efficient exon7 inclusion are shown (B and C).

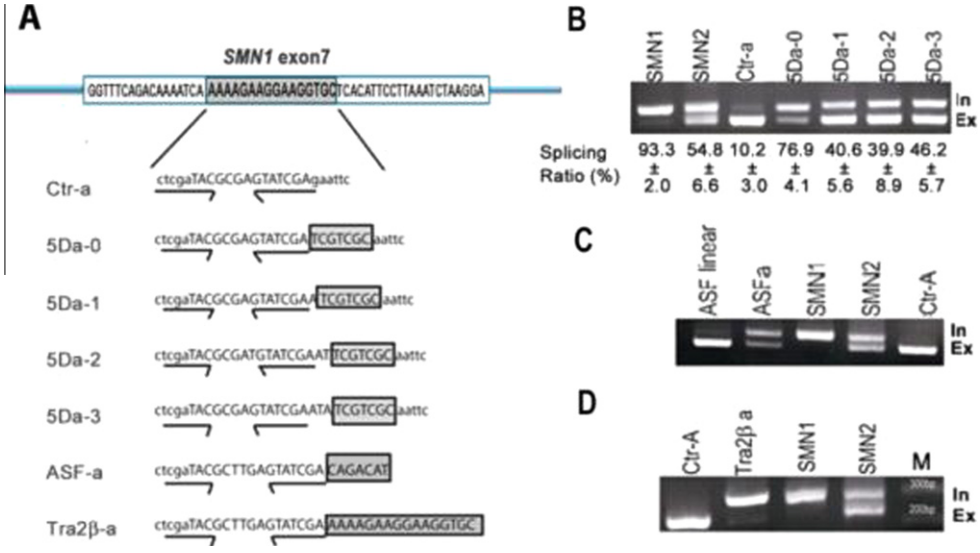

Fig. 3. The hairpin structure immediately adjacent to an ESE minimally intereferes the ESE activity. (A) The diagram of the constructs similar as in Fig. 2A. Splicing of the 5 Da (B), ASFa (C) and Tra2βa (D) series of constructs in Hela cells 24 h after transfection was assayed by RT-PCR analysis. Three sets of parallel experiments with the 5 Da constructs (B) were quantified, with the splicing ratio and standard error of the mean (S.E.M) being shown.

SMN1 exon 7, and the sequence identity was slightly mutated to alter the presence and position of the hairpin structure (Fig. 4). ESEs for SR proteins ASF, SC35, SRp40 and SRp55 were designed to cover most nucleotides in the 36-nt sequence.

In the absence of stable hairpin structures (Fig. 4B and C), these composite ESEs for all four SR proteins promoted nearly complete splicing of *SMN1* exon 7 (mulESE-SR3 and mulESE-SR4). When a stable hairpin structure (minimal free energy of −6.70 kcal/mol) is present at the 5′ → central region of the engineered sequence (mulESE-SR2), most ESEs for ASF, SRp40, SRp55 and SC35 were heavily masked by basepairing. As expected, the splicing efficiency of *SMN1* exon 7 was strongly repressed (Fig. 4B and C). The

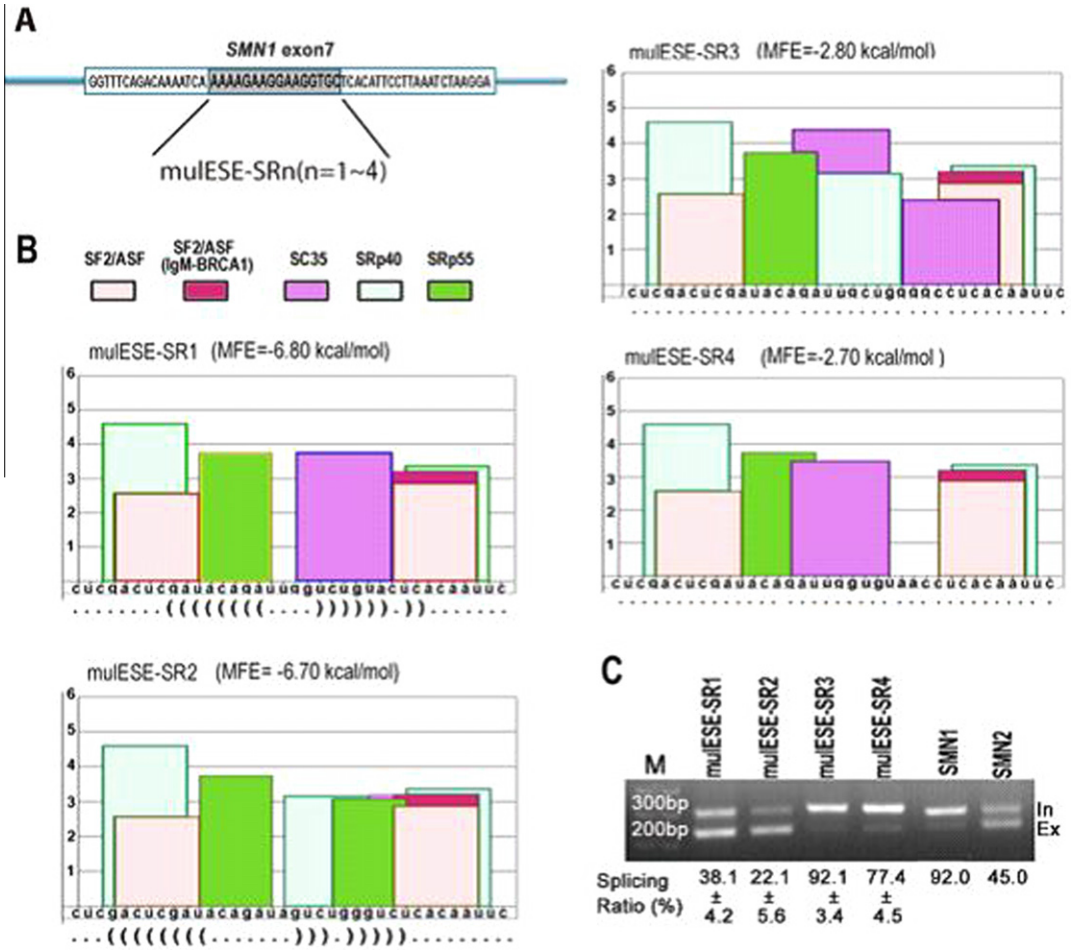

Fig. 4. A stable hairpin structure competes with the combined effect of multiple ESEs for SR proteins. (A) Diagram of individual constructs containing multiple SR protein-responsive ESEs. A 36-bp multiple-ESE (mulESE) containing sequence was cloned into the *Xho*I and *Eco*RI sites in the pZW9-*SMN1* vector. (B) Individual panels with the sequence and structural information on each construct. Shown in each panel are the exact sequence, the hairpin structure as indicated by the brackets below the sequence and the minimal free energy (MFE) predicted by RNAfold (http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi). Note that the predicted hairpin structures for the constructs muESE-SR3 and muESE-SR4 were not shown because of the low MFE values. The position (X axis) and scores of predicted ESEs (*Y* axis) for SF2/ASF, SC35, SRp40 and SRp55 (ESEfinder3.0 [44]) are shown. (C) RT-PCR analysis of each construct in transfected Hela cells. Standard errors were calculated based on three biological repeats except for the two controls that have been quantified in Fig. 3B.

remaining splicing activity might be due to an SRp40-ESE immediately downstream of the stable hairpin structure. Another downstream ASF-ESE overlapping 1-nt with the 3′ end of the hairpin may also play a role.

When the stable hairpin structure is engineered in the middle of the 36-nt sequence (mulESE-SR1, a minimum free energy of −6.80 kcal/mol), leaving structure-free ESEs for ASF and SRp40 at both ends of the hairpin structure, we observed less repressed enhancing activity (Fig. 4B and C). These results validate the structure interference of the composite ESE function. It also suggests that the downstream stem structure may protect the ESE function as does the upstream structure.

4. Discussion

4.1. Regulation of SRE function by local hairpin structures may be prevalent in mammalian cells

The stable hairpin structure folds on a microsecond time scale [35,36], much faster than the transcription process and the binding of proteins [37], suggesting that the formation of local structures on nascent mRNA is kinetically favorable in living cells. This report demonstrates that a short hairpin with a minimal stem length of 7-bp and a calculated minimal free energy (MFE) of −11.7 kcal/mol forms in human cells, which is sufficient to completely abolish the enhancer activity of a heptamer ESE nested in the stem (Fig. 1). We also show that two less stable local hairpins with calculated MFEs of −6.7 to −6.8 kcal/mol are able to strongly suppress the splicing enhancing activity of a cluster of ESEs for SR proteins. This stability requirement fits a large population of the predicted local secondary structures around the alternative splice sites [10] where SREs are enriched, indicating a prevalent biological relevance of those structures in repressing alternative splicing.

While the impact of predicted local secondary structures on ESE activities is largely consistent with the experimental data, it has been a major challenge in validating such structures in the context of the pre-mRNA in vivo by footprinting analysis. This is due to the low efficiency of footprinting chemical reagents in reaching the nuclear pre-mRNA to modify accessible nucleotides and the difficulty in obtaining enough amount of modified pre-mRNA [38]. Theoretically, ESEs in the single-stranded region of a pre-mRNA is bound by splicing factors in living cells, which would protect them from being accessed by chemical reagents just as the basepairing in the stem region does. These difficulties underscore the importance of computational prediction of secondary structures coupled with experimental determination of functional consequences as a viable

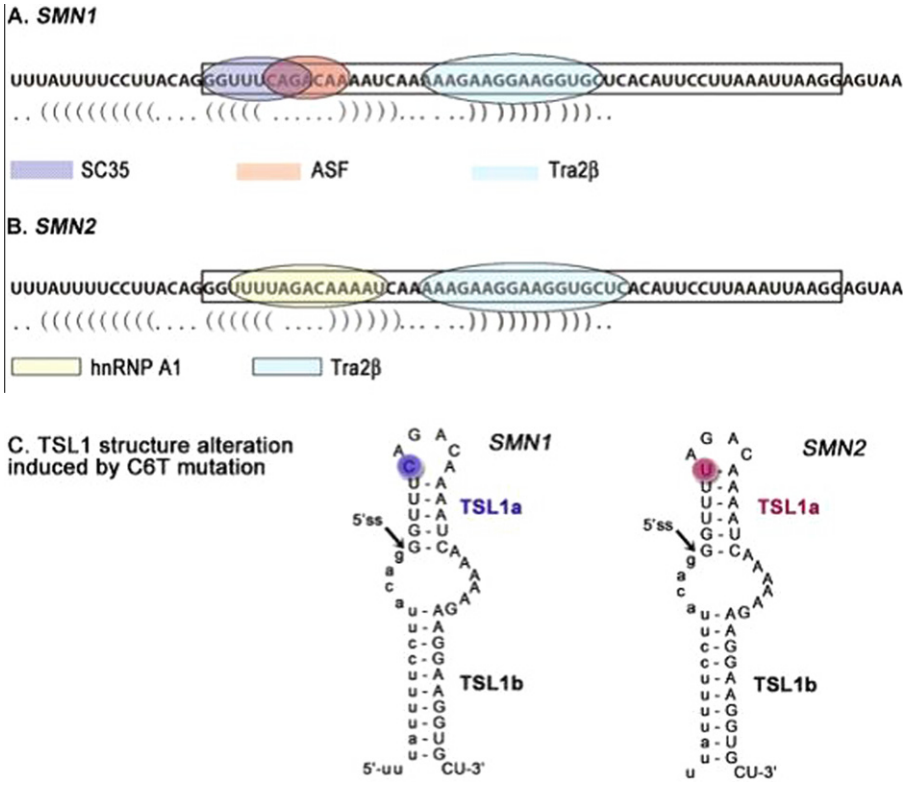

Fig. 5. The multiple effects caused by the C6T point mutation in exon 7 of the *SMN2* gene. (A and B) The exon sequences are boxed, and the basepairs in the TLS1 structure are indicated by brackets. TSL1 contains two stems (TSL1a and TSL1b) separated by 4–6 nt unpaired bulges. Binding sites for different SR proteins and HnRNP A1 protein are indicated by shaded ellipses with different colors. In *SMN1* (A), the C6 nucleotide in exon 7 is part of the binding sites for ASF/SF2 and SC35, while in *SMN2* (B), the C6T mutation removes the ESE binding sites and generates a hnRNP A1 binding (ESS). (C) Alternation of the basepairing property of TSL1 by the C6T mutation. In *SMN1*, the putative stem contains only 5 basepairs. The minimal free energy (MFE) for the formation of TSL1 in *SMN1* is −10.8 kcal/mol. An addition basepair is gained in exon 7 of SMN2, and the MFE is increased to −11.6 kcal/mol, which may contribute to exon 7 skipping in conjunction with altered interactions with SR and hnRNP proteins. The corresponding C and U in *SMN1* and *SMN2* pre-mRNAs are shaded in blue and red, respectively.

approach to complex RNA structure problems in vivo. It is widely accepted that simple hairpin structures formed locally can be reliably predicted by minimal free energy (MFE) algorithms, although *de novo* prediction of complex RNA structures involving long-range interactions is still extremely challenging [27]. The data presented here are fully consistent with the formation of the predicted local hairpin structures within a folding window of 34-nt and a MFE as high as −6.7 kcal/mol in living cells, which lends critical support to the reliability of the predicted secondary structures.

4.2. Stable upstream hairpin structures ensure the functionality of immediate downstream SREs

The advancement in developing computational algorithms for predicting SREs suggests a massive number of putative SREs [12]. However, evaluation of the regulatory activity of these SREs in a splicing reporter is problematic for two reasons. Most known SREs are within 10-nt in length and highly degenerative in sequence identity. As a result, insertion of a testing SRE into a SRE-free sequence has the potential to create some unexpected SRE(s), thus complicating the interpretation of the experimental results. In addition, insertion of an SRE may provoke the formation of some unintended secondary structures that may positively or negatively modulate the splicing activity of the SRE.

In this report, we show that placing the ASF-ESE from EPB41 gene in a hairpin loop or downstream of a stable secondary structure display a much stronger enhancer activity than in a structure-free region (Figs. 2 and 3). This observation suggests that the ASF-ESE in the structure free region may be interfered by unintend negative influences. In contrast, placing this ESE inside a hairpin loop or adjacent to a hairpin stem minimizes such influences. We also show that a CUG-rich ESE capable of forming non-canonical basepairing and base stacking interactions lost most of its enhancer activity in the hairpin loop, whereas their location after a hairpin stem guarantees its activity. Taken together, our finding suggests that an upstream stable hairpin structure can secure the regulatory activity of different classes of ESEs, which offers a general strategy for testing the splicing regulatory activity of any candidate SREs.

4.3. Natural point mutations may affect splicing by altering both the identity of cis-acting sequences and associated RNA secondary structure

Up to 25% of synonymous (in terms of amino acid coding) substitutions in exons can disrupt normal splicing [39], underscoring the contribution of silent mutations to pathogenesis via resulting splicing defects. The spinal muscular atrophy (SMA) is a neuromuscular disease characterized by degeneration of motor neurons. Approximately 95% of SMA patients lose their *SMN1* gene encoding for the survival motor neuron (SMN) protein, while *SMN2*, a nearly identical paralog of *SMN1*, remains intact in the human genome. Inefficient splicing of exon 7 (54-nt) in the *SMN2* gene is responsible for the disease phenotype because of a single C → T mutation in the 6th nucleotide (C6T) in the exon [40]. Previous study revealed that the C6T mutation abolishes the ESE consensus for two positive splicing regulators ASF/SF2 and SC35, and simultaneously create an

ESS (exonic splicing silencer) for the splicing repressor hnRNP A1 [41,42].

Interestingly, exon 7 and the adjacent intronic sequence are capable of forming two stable hairpin structures, termed TSL1 and TSL2, and the 5′ and 3′ splice sites are each located in the base-paired stem of these two hairpins. TSL2 has been proven to have negative influence on splicing [43] (Fig. 5). We note that the ESEs for the SR splicing factor ASF/SF2, SC35 and Tra2β proteins are all covered in TSL1 (Fig. 5), suggesting that disruption of the TSL1 structure may be critical for binding of these splicing factors and U2 snRNP to promote splicing of exon7. Importantly, we further note that the C6T mutation may add an additional A–U basepair in this stem, thereby increasing the thermostability of the hairpin structure (Fig. 5). We propose that, in addition to alteration of the composition of regulatory motifs (loss of ESEs and gain of ESS), the C6T mutation may also decrease the splicing efficiency of exon 7 by enforcing the local secondary structure around the 3′ splice site. While this hypothesis waits for experimental proof, our observation raises the possibility that a single point mutation in an exon may affect multiple regulatory mechanisms for the selection of the exon. Collectively, the results presented in this report emphasize the composite effect of altered cis-acting sequences and changes in associated RNA secondary structure on splicing, which may underline some dramatic functional consequences of point mutants in the human genome.

Acknowledgments

The authors are grateful to Chris Burge (MIT) and Jianhua Zhou (Arizona State University) for sending us plasmids. We are also indebted to members of the Yi Zhang lab for cooperation and discussion during this course of this investigation. This work is supported by the China 863 program (2007AA02Z112) and NSFC (30770422) to Y.Z., the China 973 program (2005CB724604) to Y.Z. and X.D.F.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.febslet.2010.09.039.

References

[1] Black, D.L. (2003) Mechanisms of alternative pre-messenger RNA splicing. Annu. Rev. Biochem. 72, 291–336.
[2] Wang, E.T. et al. (2008) Alternative isoform regulation in human tissue transcriptomes. Nature 456 (7221), 470–476.
[3] Cooper, T.A. (2005) Alternative splicing regulation impacts heart development. Cell 120 (1), 1–2.
[4] Schwerk, C. and Schulze-Osthoff, K. (2005) Regulation of apoptosis by alternative pre-mRNA splicing. Mol. Cell 19 (1), 1–13.
[5] Li, Q., Lee, J.A. and Black, D.L. (2007) Neuronal regulation of alternative pre-mRNA splicing. Nat. Rev. Neurosci. 8 (11), 819–831.
[6] Biamonti, G. and Caceres, J.F. (2009) Cellular stress and RNA splicing. Trends Biochem. Sci. 34 (3), 146–153.
[7] Chen, M. and Manley, J.L. (2009) Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. Nat. Rev. Mol. Cell Biol. 10 (11), 741–754.
[8] Buratti, E. and Baralle, F.E. (2004) Influence of RNA secondary structure on the pre-mRNA splicing process. Mol. Cell Biol. 24 (24), 10505–10514.
[10] Shepard, P.J. and Hertel, K.J. (2008) Conserved RNA secondary structures promote alternative splicing. RNA 14 (8), 1463–1469.
[11] Warf, M.B. and Berglund, J.A. (2010) Role of RNA structure in regulating pre-mRNA splicing. Trends Biochem. Sci. 35 (3), 168–178.
[12] Fairbrother, W.G. et al. (2002) Predictive identification of exonic splicing enhancers in human genes. Science 297 (5583), 1007–1013.
[13] Wang, Z. et al. (2004) Systematic identification and analysis of exonic splicing silencers. Cell 119 (6), 831–845.
[14] Fu, X.D. (2004) Towards a splicing code. Cell 119 (6), 736–738.
[15] Tejedor, J.R. and Valcarcel, J. (2010) Gene regulation: breaking the second genetic code. Nature 465 (7294), 45–46.
[16] Barash, Y. et al. (2010) Deciphering the splicing code. Nature 465 (7294), 53–59.
[17] Wang, G.S. and Cooper, T.A. (2007) Splicing in disease: disruption of the splicing code and the decoding machinery. Nat. Rev. Genet. 8 (10), 749–761.
[18] Xue, Y. et al. (2009) Genome-wide analysis of PTB-RNA interactions reveals a strategy used by the general splicing repressor to modulate exon inclusion or skipping. Mol. Cell 36 (6), 996–1006.
[19] Yeo, G.W. et al. (2009) An RNA code for the FOX2 splicing regulator revealed by mapping RNA-protein interactions in stem cells. Nat. Struct. Mol. Biol. 16 (2), 130–137.
[20] Licatalosi, D.D. et al. (2008) HITS-CLIP yields genome-wide insights into brain alternative RNA processing. Nature 456 (7221), 464–469.
[21] Buratti, E. et al. (2004) RNA folding affects the recruitment of SR proteins by mouse and human polypurinic enhancer elements in the fibronectin EDA exon. Mol. Cell Biol. 24 (3), 1387–1400.
[22] Muro, A.F. et al. (1999) Regulation of fibronectin EDA exon alternative splicing: possible role of RNA secondary structure for enhancer display. Mol. Cell Biol. 19 (4), 2657–2671.
[23] Hiller, M. (2007) Pre-mRNA secondary structures influence exon recognition. PLoS Genet. 3 (11), e204.
[24] Liu, H. and Naismith, J.H. (2008) An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol. BMC Biotechnol. 8, 91.
[25] Eperon, L.P. et al. (1988) Effects of RNA secondary structure on alternative splicing of pre-mRNA: is folding limited to a region behind the transcribing RNA polymerase? Cell 54 (3), 393–401.
[26] Schroeder, R. et al. (2002) RNA folding in vivo. Curr. Opin. Struct. Biol. 12 (3), 296–300.
[27] Shapiro, B.A. et al. (2007) Bridging the gap in RNA structure prediction. Curr. Opin. Struct. Biol. 17 (2), 157–165.
[28] Mahen, E.M. et al. (2010) mRNA secondary structures fold sequentially but exchange rapidly in vivo. PLoS Biol. 8 (2), e1000307.
[29] Yang, G. et al. (2005) An erythroid differentiation-specific splicing switch in protein 4.1R mediated by the interaction of SF2/ASF with an exonic splicing enhancer. Blood 105 (5), 2146–2153.
[30] Shen, L.X., Cai, Z. and Tinoco Jr., I. (1995) RNA structure at high resolution. FASEB J. 9 (11), 1023–1033.
[31] Baumruk, V. et al. (2001) Comparison between CUUG and UUCG tetraloops: thermodynamic stability and structural features analyzed by UV absorption and vibrational spectroscopy. Nucleic Acids Res. 29 (19), 4089–4096.
[32] Denisov, A.Y. et al. (2003) A novel RNA motif based on the structure of unusually stable 2′, 5′-linked r(UUCG) loops. J. Am. Chem. Soc. 125 (38), 11525–11531.
[33] Desmet, F.O. et al. (2009) Human Splicing Finder: an online bioinformatics tool to predict splicing signals. Nucleic Acids Res. 37 (9), e67.
[34] Wang, J. et al. (2005) Distribution of SR protein exonic splicing enhancer motifs in human protein-coding genes. Nucleic Acids Res. 33 (16), 5053–5062.
[35] Gralla, J. and Crothers, D.M. (1973) Free energy of imperfect nucleic acid helices II. Small hairpin loops. J. Mol. Biol. 73 (4), 497–511.
[36] Ravetch, J., Gralla, J. and Crothers, D.M. (1974) Thermodynamic and kinetic properties of short RNA helices: the oligomer sequence AnGCUn. Nucleic Acids Res. 1 (1), 109–127.
[37] Darzacq, X. et al. (2007) In vivo dynamics of RNA polymerase II transcription. Nat. Struct. Mol. Biol. 14 (9), 796–806.
[38] Liebega, A. and Waldsich, C. (2009) Probing RNA structure within living cells. Methods Enzymol 468, 219–318.
[39] Pagani, F., Raponi, M. and Baralle, F.E. (2005) Synonymous mutations in CFTR exon 12 affect splicing and are not neutral in evolution. Proc. Natl. Acad. Sci. USA 102 (18), 6368–6372.
[40] Cooper, T.A., Wan, L. and Dreyfuss, G. (2009) RNA and disease. Cell 136 (4), 777–793.
[41] Cartegni, L. and Krainer, A.R. (2002) Disruption of an SF2/ASF-dependent exonic splicing enhancer in SMN2 causes spinal muscular atrophy in the absence of SMN1. Nat. Genet. 30 (4), 377–384.
[42] Kashima, T. and Manley, J.L. (2003) A negative element in SMN2 exon 7 inhibits splicing in spinal muscular atrophy. Nat. Genet. 34 (4), 460–463.
[43] Singh, N.N., Singh, R.N. and Androphy, E.J. (2007) Modulating role of RNA structure in alternative splicing of a critical exon in the spinal muscular atrophy genes. Nucleic Acids Res. 35 (2), 371–389.
[44] Cartegni, L. et al. (2003) ESEfinder: a web resource to identify exonic splicing enhancers. Nucleic Acids Res. 31 (13), 3568–3571.